Real-Time Operational Planning for the U.S. Air Traffic System

John M. Mulvey
Stavros A. Zenios

Report EES-86-5
Civil Engineering Department
Engineering-Management Systems
Princeton University
Princeton, NJ 08544

ABSTRACT

This report describes a planning model for the continental U. S. air traffic system. The basic approach employs the dual objectives of monitoring collision risks while minimizing transportation costs. Due to the model's special structure-- a network graph -- extremely efficient nonlinear algorithms are available for solving problems in this class. Test problems from the Indianapolis control sector are solved with a CRAY X-MP/24 supercomputer. Despite these results, further work is needed to develop a practical system, given current hardware/software technology. Suggestions are made for combining advances in artificial intelligence and mathematical modeling.

1. Introduction

Computational and information technologies have gotten to the stage where large-scale *operational* planning models are rapidly gaining acceptance. Some common examples are the following applications. Trucking companies are beginning to employ stochastic networks for scheduling and pricing their "less than truckload" shipments in response to deregulation. Airlines routinely assign pilots/crews to routes so as to minimize costs. Railroads manage their freight-cars by means of elaborate information-decision systems. In each of these cases, large amounts of data are assembled in the context of a mathematical, real time planning model. Decisions are rendered in conjunction with this model, using the expert judgment of the decision-maker [DM]. It is important to recognize that the process takes advantage of the best attributes of both the human *and* the computer. Each is essential.

One of the largest operational planning problems involves the real-time control of the United States air traffic system. On average, in 1984 over 17,000 flights per day traveled the high level (above 29,000

This research was funded in part by NSF grant DCR-8401098 and NASA grant NAG-1-520.

(NASA-CR-181046) REAL-TIME OPERATIONAL PLANNING FOR THE US AIR TRAFFIC SYSTEM (Princeton Univ.) 21 p Avail: NTIS
N87-70426
Unclas
00/04 0079396

feet) jet routes. It is anticipated that traffic will increase by 40% over the next 5 years. In response to this build up, the U. S. Federal Aviation Administration (FAA) has put together the National Airspace System Plan [NASP][1]. Its goal is to coordinate the design and implementation of appropriate new hardware/software for the air traffic system. The FAA will spend over $11 billion in conjunction with the NASP.

While there are several possible modes of operations for integrated risk/cost planning, this report will concentrate on an operational model [OP]. This particular model has interest for several reasons. First, the problem is extremely large, even by today's supercomputer standards, due to the complexity of the feasible routes and possible time delays. It is projected that an 8-period (e.g., eight 1-hour time steps) regional model will consist of approximately 15,000 equations and 270,000 nonlinear variables. To be practical for operational planning, a problem of this magnitude must be solved in less than 30 minutes of elapsed time.

Second, there are a number of conflicting objectives to be considered: a) systemwide risk, b) aggregate transportation costs, c) individual carrier costs and delays, and so on. Tradeoffs in these criteria will undoubtedly require a systematic method for assessing the implicit marginal rates of substitution (MRS) between the conflicting criteria. Also note that MRS's will depend upon the level of satisfaction for each particular criterion affected.

Finally, stochastic elements must be considered along with combinatorial aspects. The weather and other random affects will often cause the forecasted time of flight to be inaccurate. Since the model consists of multiple time periods, a multi-scenario planning procedure would provide an ideal framework; however, the problem's enormous size precludes this approach in the foreseeable future. Another strategy is to operate under a rolling horizon in which decisions beyond the current period are only tentative and are updated as time unfolds. This approach has the ability to adapt as conditions warrant -- a form of stochastic programming.

2. Basic Planning Model

Solving large scale problems requires taking advantage of the problem's special structure. This strategy has been particularly successful in the case of network optimization algorithms. See references [2,9,10] for examples. Today, a linear network in excess of 20,000 nodes (equations) and 2,000,000 arcs (variables) can be solved routinely.

The initial operational planning model is defined below as a bicriteria problem:

[OP] *Minimize* $\alpha_0 \cdot \Omega(\bar{x}) + \alpha_1 \cdot T(\bar{x})$

Subject to :

$$\sum_{r \in R_f^t} \sum_{t \in T_f} \sum_{f \in F} \Gamma_{f,r,t}^{g,s} \cdot x_{f,r}^t \leq L_{g,s} \quad \text{for } all \ g \in G \ , s \in T \tag{1}$$

$$x_{f,r}^t = \begin{cases} 1 & flight\ f\ takes\ route\ r, departs\ at\ time\ t \\ 0 & otherwise \end{cases} \tag{2}$$

$$\sum_{r \in R_f^t} \sum_{t \in T_f} x_{f,r}^t = 1 \quad \text{for } all\ f \in F \tag{3}$$

$$T(\bar{x}) \equiv \sum_{r \in R_f^t} \sum_{t \in T_f} \sum_{f \in F} c_{f,r}^t \cdot x_{f,r}^t \tag{4}$$

where :

T $(\bar{x})$: transportation costs

$\Omega(\bar{x})$: system risks

flights : f ∈ F

time periods : t ∈ T = {1,2,3, ... ,p}

eligible departing times for flight f : $\{T_f\}$

routes : r ∈ R

eligible routes for flight f, departing at time t : $\{R_f^t\}$

cost for flight f, departing at time t, taking route r : $c_{f,r}^t$

geographical regions : g ∈ G

limit on number of airplanes occupying region g, during time period s : $L_{g,s}$

incidence matrix for flight f, departing at time t, taking route r, in region g at time s : $\Gamma_{f,r,t}^{g,s}$

The two parameters in the objective function -- α_0 and α_1 -- represent goal programming weights for the system risk $\Omega(\bar{x})$ and the transportation cost T($\bar{x}$). The transportation cost function includes both flight travel cost and delay costs, either on the ground or in the air. Typically, the risk function $\Omega(\bar{x})$ will be non-linear, nonseparable and sparse.

The planning model has been defined in a very general setting, using arbitrary geographical regions g $\in$ G and multiple time periods t $\in$ T. In fact the regions may overlap or depict geography as detailed as individual runways. Some examples of regions are the following : departing airports, en-route control sectors and destination airports. A special case of [OP] is the "flow control" problem[6,4].

Flights, f $\in$ F , are also defined in a very general manner. They may represent the usual { airport a to airport b } link, or they may represent { point c to airport b } or any general path across the United States. Here, point c depicts an intermediate air-space location between airports a and b.

The constraints (3) ensure that every flight is scheduled to a single route. Constraints (1) are based on airspace and airport limitations; these are defined as limits on the amount of traffic which can be safely monitored in various regions (or sectors). These rules are dependent, of course, on the type of traffic which will be using the facilities, anticipated weather and other parameters.

Mention should be made of the incidence matrix Γ. This array identifies, for each x-variable, the time periods (s) and the regions (g) which will be traversed by flight $x^t_{f\,r}$. Remember that $x^t_{f\,r}$ represent flight f $\in$ F, departure t $\in$ T_f and route r $\in$ R^t_f.

It should be noted that model [OP] can be interpreted in several ways:

(1) transhipment network with multicommodity side constraints and integer variables.

(2) integer set partitioning problem with capacity constraints, or

(3) general integer program (IP).

Because of the model's ultra-large size, it is imperative to develop a solution algorithm which is extremely efficient. Not only must the model be solved in a real-time environment but also it must take into account possible uncertainties, for instance in travel time.

In a previous report [11] we have shown that a network approach to solving model [OP] is feasible. Section 5 reviews the nonlinear network algorithm for solving [OP] and reports computational results. Figure 1 depicts a small sample problem consisting of 5 airports and 10 decision variables. In this example, risk is handled by means of a congestion function which limits the number of flights in any well defined airspace, over a specified time period. If congestion is predicted, the model considers two courses of action: either flights are delayed in time, or they are assigned to a new geographical path. The model minimizes total transportation costs subject to congestion restrictions (at the airports and in the air).

An important feature is the model's ability to provide "optimal" solutions. Flight delays, especially those occurring in the air, are expensive due to the costs of fuel, salaries for the crew, ill-will invoked from the passengers, and so on. A 1% gain in operating efficiency translates into multi-million dollar savings in fuel costs alone. Non optimizing heuristic algorithms are difficult to justify when an optimization framework is available and when the opportunity costs are so high. In conjunction with cost minimizing, the [OP] model will provide a safer environment due to improved flight management. Peak congested regions will be reduced, since constraints (2) limit the maximum number of flights in a given time-airspace. Stresses on air traffic controllers will be reduced, and the workloads will be distributed more evenly.

3. Model Uses

The primary application of the proposed planning model is to identify any anticipated instances of congestion and to propose alternative actions. Note that the model considers the limitations of the airports as well as other sources of possible congestions in handling traffic patterns. As such, the flow-control management problem depicts a special case of model OP. "Flow control" delays aircraft on the ground when congestion is anticipated at the destination airport[6,4]. (Currently, the nine busiest airports are monitored for congestion as part of this delay system.)

Figure 2 illustrates a highly simplified example. Here flights 101 and 102 are scheduled to traverse sector xyz during period t. However, flight 103 and 104 have also requested the same sector during the same time period. Two types of decisions can be made in order to keep the total number of flights below the pre-specified maximum --in this case three: either flights must be delayed or a different path must be taken by at least one of the airplanes. Whatever is chosen by the model must also be assessed by the air traffic controller. Yet the decision will affect other neighboring flights, in a cascading manner, and these must be considered by the model.

As an operational tool, model [OP] provides an ideal framework for collecting and monitoring timely flight information on a national basis. At present, it it extremely difficult to evaluate the network aspects of

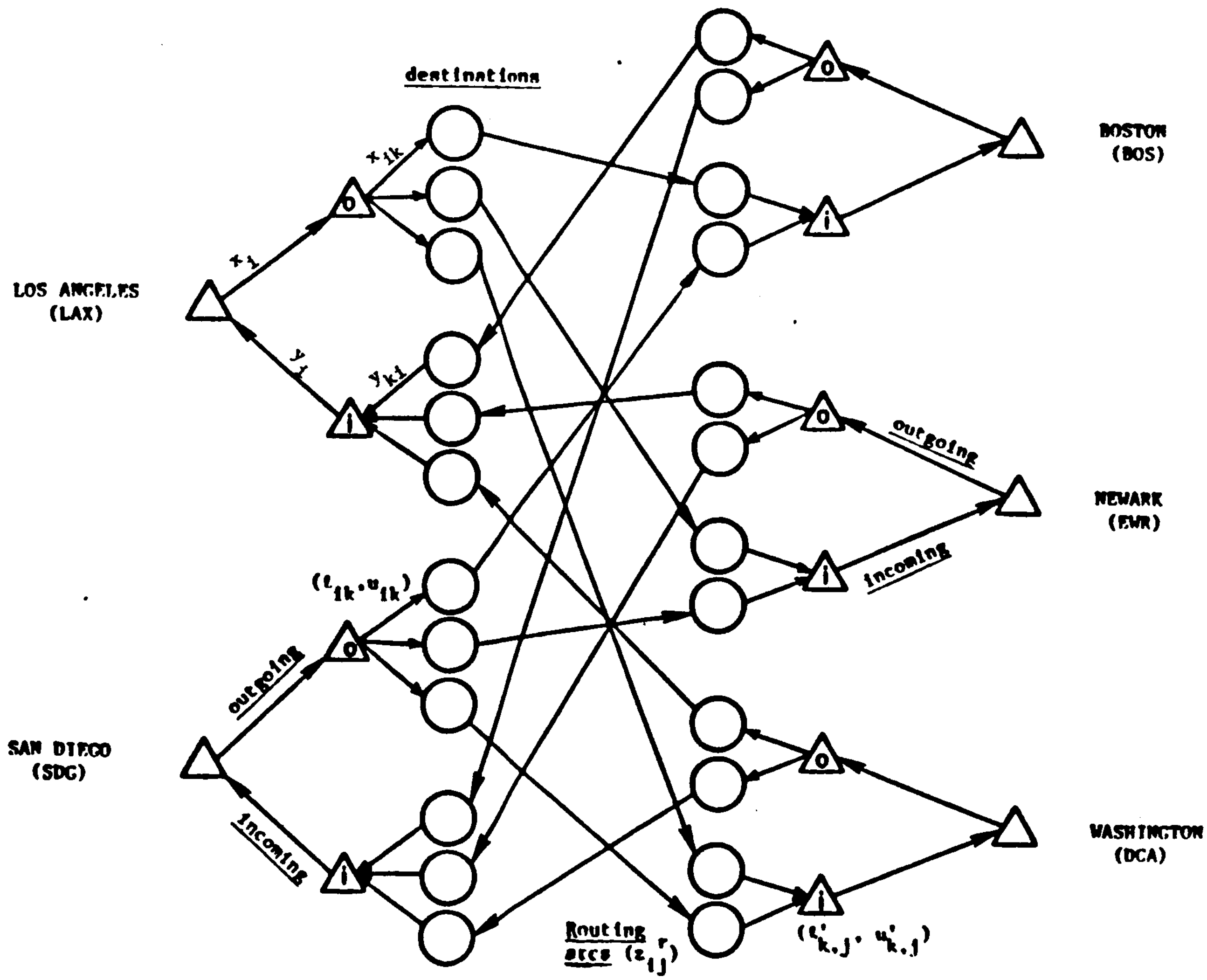

Figure 1: A Sample Air-traffic Control Model

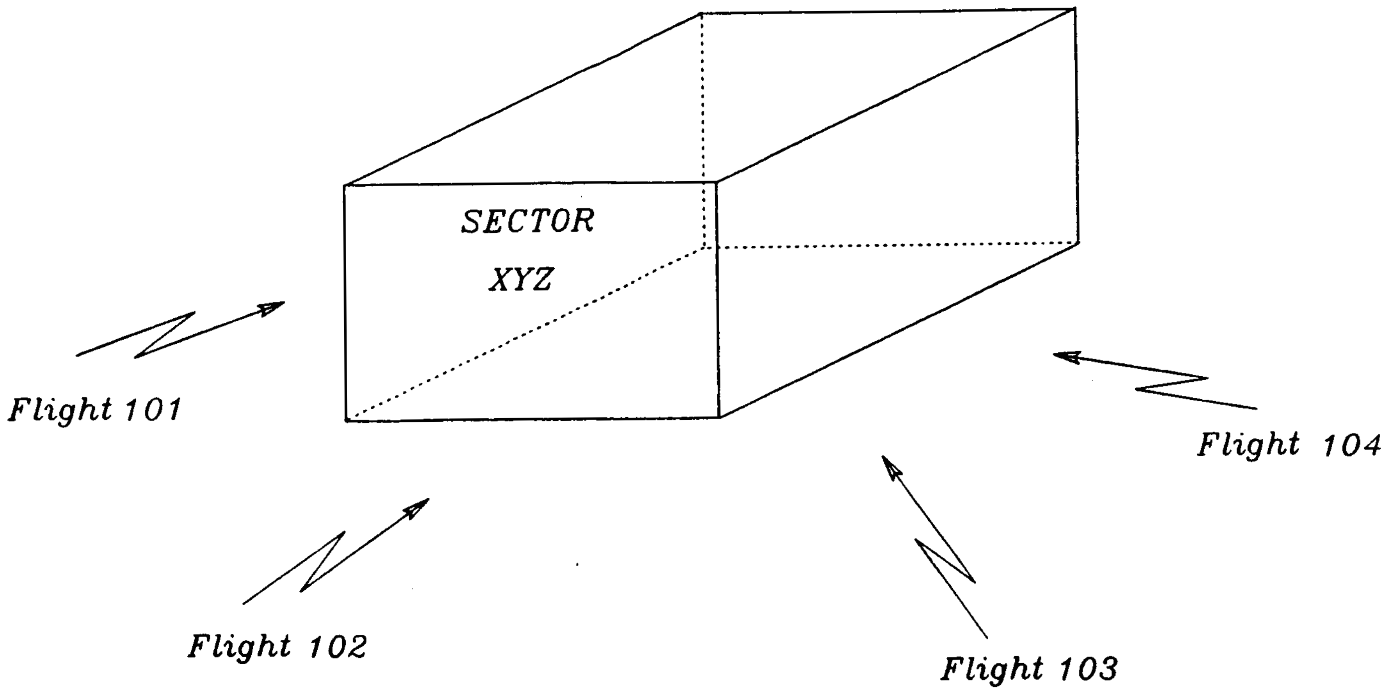

Figure 2: A Simple Example of Air-traffic Control Model

risks, except through the mechanism of counting actual collisions or "close encounters", or by counting the number of controler interventions required to avoid such encounters. The network database could be used within the context of a comprehensive risk evaluation system. The result would be a more accurate evaluator of network risk and would compliment the more traditional approaches (Odoni and Endoh[12]) which are based on the analysis of individual micro-events. As the forecasted increases in traffic occur during the next few years, new pressures will be placed on the (systemwide) command-and-control aspect of U.S. air traffic.

One of the current difficulties encountered in the flow-control procedures mentioned previously is perceived inequality among individual air-carriers. Carrier xyz believes that they are receiving an unfair percentage of the delays. And perhaps, the airline is correct in their perception. Not only would model OP measure the degree (and kind) of delays for each air-carrier but also the model could be easily expanded to include equity considerations, through a modified objective function or additional linear constraints.

Restrictions would take the following form:

$$\sum_{r\in R_j^t}\sum_{t\in T_f}\sum_{f\in F_c} d_{f,r}^t \cdot x_{f,r}^t - (AVE) - \varepsilon_c^+ + \varepsilon_c^- = 0\,, \quad \text{for } all\ carriers\ c \in C$$

where

(AVE) is the mean delay for the air-traffic system, callibrated for the carriers.

$d_{f,r}^t$: total ground (air) delay for flight f, departing time t, route r.

$\{F_c\}$: all flights for carrier c

ε_c^+ , ε_c^- : positive and negative derivations for carrier c

Another method for handling inequitable delays is to invoke a side payment plan, whereby under-delayed carriers compensate over-delayed carriers. While in theory, a loser compensation plan is more efficient than the method in which extra constraints are imposed, there are potential problems. First, the system is subject to abuse. Second, the system requires a bureaucracy for monitoring and maintaining itself. Third, antitrust laws must be dealt with. Further research is needed to resolve these policy issues.

As experience is gained with the basic OP model features and options can be added. For example, many people have argued that a deregulated air traffic system, in which airplanes are allowed to take any path, would be an improvement over the fixed route structure which is largely in place today. Model OP could be easily expanded to include variable routes as long as the ensuring model size is within the range of practical solution. The design of efficient solution algorithms is essential in this extension.

The interface between the human decision maker and a model such as OP will require extensive research so that the coordinated action is virtually foolproof. Information must be presented in a form which is suitable (probably graphically) for the user; he must be able to respond in a convenient and unambiguous manner. This will require novel ideas from artificial intelligence and operations research.

4. Tactical Considerations

Selecting an acceptable level of congestion for a particular airspace (i. e. the $L_{g,s}$ and α_0 coefficients) is a complicated task. The decision has several dimensions. On the aggregate level, congestion can be traced out against total travel costs as shown in figure 3 as an efficient frontier. Or the congestion

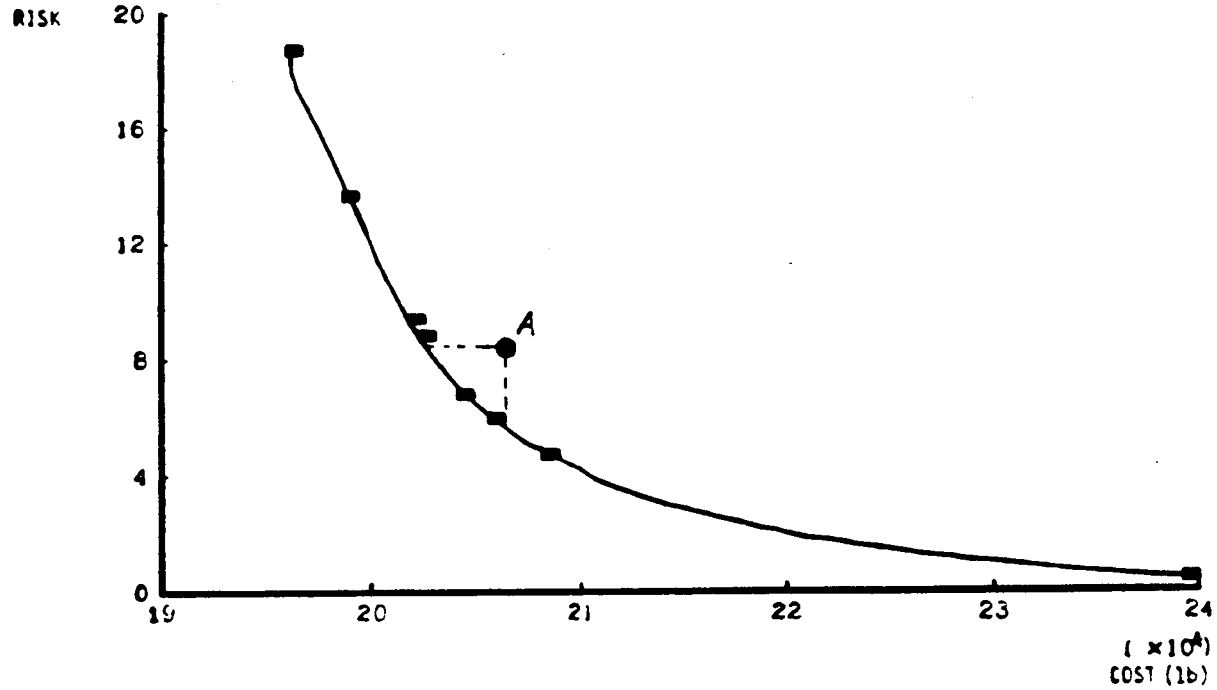

Figure 3: Efficient Frontier for Risk/Cost Tradeoff

criteria could be measured on an average basis, rather than a worst case basis. Regardless of these issues, the parameters will be influenced by the details of geometry and the types of aircraft which fly through. The OP model provides enough generality for handling these factors in a systematic fashion.

Other tactical decisions regarding the air system can be analyzed. For instance, the minimum spacing required between aircraft will affect the number and type of decision variables. Given a proposed modification, the OP model can be run before and after; differences in performance provide an estimate of the impact of the proposed alteration. Figure 4 shows a hypothetical efficient frontier for two cases. The two criteria in this situation are maximum congestion and total costs, as indicated along the two axes. This analysis provides a systematic framework for deciding if the proposed modification is sensible. However, one must be careful. Again turning to Figure 4, we see that the original system is more desirable when the maximum level of congestion lies below the threshold value c_0. Conversely, the new system appears more desirable when the threshold value exceeds c_0. The decision hinges on the acceptable level of maximum congestion.

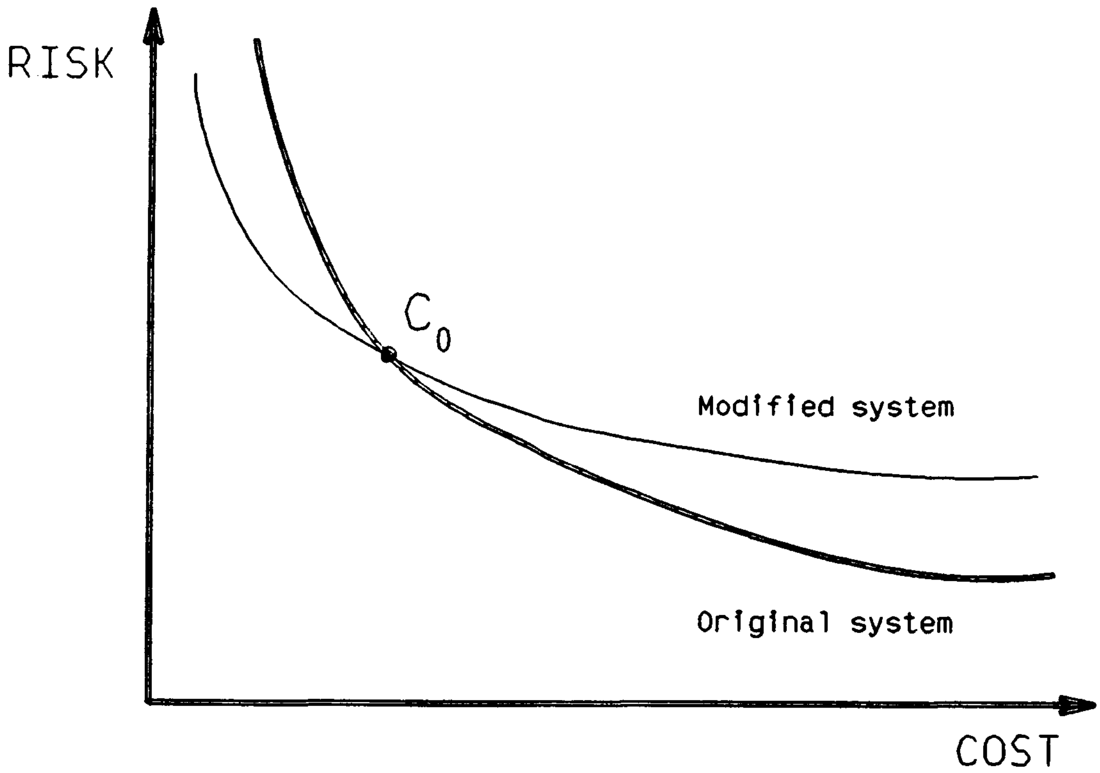

Figure 4: Comparing two Operational Scenarios

5. Algorithmic Issues

This section takes up issues which are pertinent to the solution of ultra-large network problems. To accomplish this task will require the interplay between two diverse disciplines. On the one hand specialized algorithms are required that will capitalize on the special structure of the network models. On a different front we should consider accessing a supercomputer, that has enough computational power and memory size to handle the ultra large problems arising from the OP model. Today, all supercomputer designs employ some form of multi-processing -- either vector or distributed. Hence, an important effort will be to adapt the network algorithms for these machines. Research is already under way for streamlining the network specialized algorithms for the environment of vector computers, see for example Zenios and

Mulvey[14]. In addition developments are taking place in designing numerical algorithms that can be processed efficiently on massively distributed systems; see Bertsekas and El Baz[3] or Zenios and Mulvey[13]. We describe in this section a specialized algorithm for solving problems of the form [OP]. Through computational testing the efficiency of this algorithm on a wide range of computer systems will be established.

5.1. Truncated Newton Algorithm

In a manner similar to most nonlinear programming procedures, each truncated Newton (TN) iteration consists of two stages : (1) a search direction routine, and (2) a step length routine. Table 1 depicts the overall flow. The search direction must fulfill certain essential features so that the overall algorithm will converge and so that performance efficiencies are attained. First, the direction must both maintain feasibility and point downhill (in a minimization context). Defining the search direction as $\bar{p}^k$ and given a feasible point $\bar{x}^k$ at the k^{th} iteration, the usual Newton method for calculating $\bar{p}^k$ would solve the following quadratic programming problem :

[QP] *Minimize* $\frac{1}{2}(\bar{p}^k)^t\, G(\bar{x}^k)\,\bar{p}^k + g(\bar{x}^k)^t\,\bar{p}^k$

Subject to :

$$\bar{A}\,.\,\bar{p}^k = 0$$
$$p_j^k \geq 0 \text{ if } x_j^k = l_j$$
$$p_j^k \leq 0 \text{ if } x_j^k = u_j$$

<u>where</u>

$g(\bar{x}^k)$ gradient at $\bar{x}^k$

$G(\bar{x}^k)$ Hessian at $\bar{x}^k$

By restricting our attention to a special projected matrix Z, whose columns form a basis for the null space of A, i.e. $A \cdot Z = 0$, the problem [QP] can be solved using the following two formulae :

$$(Z^t\, G\, Z)\cdot \bar{p}_s^k = Z^t\, \bar{g}^k \qquad [5.1]$$
$$\bar{p}^k = Z^t\, \bar{p}_s^k$$

<u>where</u>

$$Z \equiv \begin{array}{cc} \begin{bmatrix} -B^{-1}S \\ S \\ N \end{bmatrix} & \begin{array}{c} m \\ s \\ n-s-m \end{array} \end{array}$$

G semi-positive approximation to the Hessian at point $\bar{x}^k$

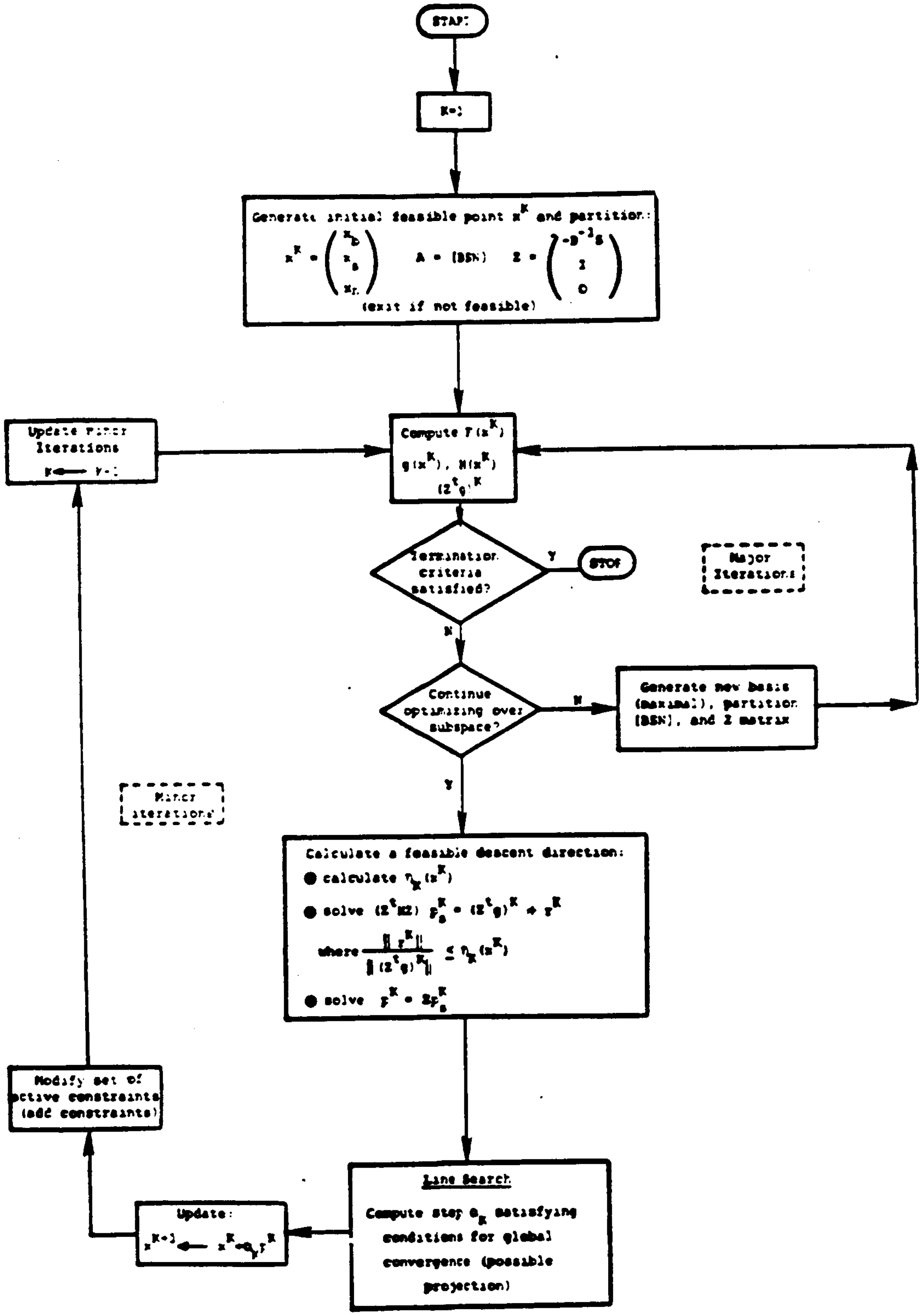

Table 1 : The Truncated Newton Algorithm

$\bar{g}^k$ gradient of objective function at point $\bar{x}^k$

and where the decision variables have been partitioned into three sets :

$$\begin{aligned} \bar{x} &= [x_b \mid x_s \mid x_n] \\ A &= [B \mid S \mid N] \\ g(\bar{x}) &= [g_b \mid g_s \mid g_n] \\ \bar{p}^k &= [p_b \mid p_s \mid p_n] \end{aligned}$$

The benefits of the Newton direction $\bar{p}^k$ are greatest in the neighborhood of a solution; however it is expensive to calculate the solution of equation [5.1]. In response, we adjust in a dynamic fashion the degree of accuracy of solving [QP]. A forcing sequence $\{\eta^k\} \rightarrow 0$ is employed in this regard. Accuracy is defined according to the relative residual in equation [5.1], $\frac{||r^k||}{||Z^t \bar{g}^k||}$, in which $r^k \equiv (Z^t\ G\ Z)\,.\,\bar{p}_s^k + Z^t\ \bar{g}^k$ and $||.||$ is a vector norm in R^n. The minor iteration (see Table 1) continues only until the required accuracy is attained. Thus, $\frac{||r^k||}{||Z^t \bar{g}^k||} \le \eta^k$ defines the termination criteria for the minor iterations.

When the algorithm is far from the solution the reduced gradient -- $||Z^t\ \bar{g}^k||$ -- is large and little work is required to locate a direction satisfying the acceptance criteria. Only the basic and super basic variables are optimized. If one of these variables hits a bound, the constraining variable is transferred into the set of nonbasics $[x_n]$. As $||Z^t\ \bar{g}^k||$ is reduced the acceptance criteria becomes more restrictive and the current solution to the direction finding problem lies closer to the Newton direction.

At this point, the nonbasic variables must be tested for optimality. First order estimates for the Lagrange multipliers are computed as follows :

$$\begin{aligned} \bar{\mu}_b^t &= \bar{g}_b^t\,.\,B^{-1} \\ \bar{\mu}_n^t &= [\bar{g}_n^t - \bar{\mu}_b^t\,N] \\ \bar{\mu}_s^t &= \bar{\mu}_b^t\ S \end{aligned}$$

In this environment non-basic variables that reduce the objective function when moving away from their bounds (i.e., if $\mu_n^j < 0$ and $x_n^j = u_n^j$, or if $\mu_n^j > 0$ and $x_n^j = l_n^j$) along with free non-basic variables (i.e., $l_n^j < x_n^j < u_n^j$) are called eligible. Eligible variables are transferred to the superbasic set $[x_s]$ in conjunction with a maximal basis[7], and the TN algorithm continues the next major iteration with the new partition.

A sizable portion of the algorithm's execution time involves computing the search direction $\bar{p}_s^k$. While the Truncated-Newton method can use any iterative method for solving equation [5.1], we have chosen the linear conjugate-gradient [CG] method. Although the reduced Hessian matrix $Z^t GZ$ is typically dense, the product required by [CG], $(Z^t GZ)\,\bar{v}$, is easily computable due to the sparsity of the large-scale components. The success of the conjugate-gradient method depends upon locating a "good" search direction in a small number of iterations. Thus, preconditioning the reduced Hessian by the matrix P is important so as to reduce the number of CG iterations. Whereas the usual initial element of the CG sequence is $\bar{g}_s^k$, the vector $P\ .\ \bar{g}_s^k$ becomes an initial element when preconditioning, where P is a positive-definite matrix. See [2] for further details.

6. Data Requirements and the Indianapolis Control Sector

One of the most critical aspects in modeling real world systems is the ability to collect the required data in a timely fashion. In the case of the [OP] models, data may be classified in two categories : static and dynamic. By static we refer to data that do not change over a long period of time (e.g. airport locations) and by dynamic we mean data that change with time (e.g. flights scheduled), or with technological innovation (e.g. aircraft fuel burn rates, navigation systems, flight management procedures). For the model to be useful, the input data must readily available. A key component of the comprehensive National Airspace System Plan is a centralized data-base of aircrafts scheduled for, or actually flying the high altitude jet routes[8]. This data-base can serve on a real-time basis for updating the data required for [OPT].

For the prototype model developed the following categories of data were required :

(I) Airports information

(II) Flights information

(III) Fuel Burn data

Table 3 provides more details. Some data, like the airports coordinates, were available through sources used in the past by FAA, while other data had to be collected for the network model.

The following sources were employed:

(I) International Official Airlines Guide (IOAG) tape, providing information about the airports.

(II) Flight Progress Strip data collected by the Control Center, providing flights information .

(III) Fuel Burn Model developed by the FAA, providing data about fuel burn rates for different types of aircrafts.

• **Airports Information** 1. Airport ID code 2. Geographical coordinates
• **Flight Information** 1. Flight ID 2. Origin airport 3. Destination airport 4. Cruise altitude on entering target sector 5. Cruise altitude on exit from the target sector 6. Time flight enters the target sector 7. Time flight exits from the target sector 8. Flight Hemi code defining legitimate cruise altitudes
• **Fuel Burn Data** 1. Aircraft type 2. Fuel burn rate per hour for every legitimate cruise altitude 3. Fuel burn rate per nautical mile for every legitimate cruise altitude

Table 3 : **Model data requirements**

A network model was built for the airspace controlled by a sector of the Indianapolis center. The purpose of the model is to serve as a prototype to illustrate the use of the optimization algorithms described earlier as well as the feasibility of the proposed model. Data were collected for a high traffic period on January 9, 1985, in which a total of 185 aircrafts crossed the sector over a 6 hour period. The duration of flight through the sector ranged from 4 to 23 minutes. Five distinct cruise altitudes above 29000 feet (FL2900) were selected by the planes. The following provisions were made in the model :

(I) Allow for delays at the origin airport, up to three 10-minute intervals and similar delays at the destination airports. This time grid can be made finer by considering a larger number of progressively smaller delay intervals (e.g. six 5-minute intervals). The added accuracy will be balanced by the larger network that has to be solved.

(II) Allow for every plane to follow one or two alternative cruise altitudes besides the one currently followed. Choice was restricted to the cruise altitudes one level above and one level below the primary altitude. Again, this restriction can be relaxed at the expense of generating larger network models -- the aircrafts could be instructed to follow any one of the four or five legitimate cruise altitudes, specified for the particular flight.

The model includes the dual objective of assessing risk and cost , as proposed in an earlier section. For more details on modeling this particular control sector refer to the paper by Mulvey and Zenios[11]. The developed model was used as a prototype on which the developed algorithmic tools were tested. The availability of realistic data also assists in validating the proposed modeling framework. By varying the relative weights on the transportation and risk functions in a systematic way the risk/cost efficient frontier was traced (Figure 3). Again, the efficient frontier is not meant to serve as a direct way of comparing risk with cost. Instead it guides one in evaluating alternative modes of operation of the air-traffic control system, as generated by the model, or with currently followed procedures. The major advantage of this methodology is that it generates a sequence of alternatives that are efficient; i.e. both risk and cost measures cannot be improved simultaneously. This is easier to understand if we notice the location of point A in Figure 3 - this point was obtained by solving the optimization problem inexactly. From point A we may move to a series of alternative solutions for which the system is better off, both with respect transportation cost and risk.

To study the effect of airplane congestion, we developed a histogram of all planes flying at a particular altitude, during the ten time intervals of interest. Figure 5 summarizes the results for three particular altitudes, before and after the optimization model was used. Note that, as expected, planes were diverted from a highly congested altitude (35000 ft) to less congested routes (31000 ft and 39000 ft). This result was obtained with relative weights 0.5/0.5 on both risk and transportation costs.

Problem	NLPNETG Solution times (sec)		
	IBM 3081	VAX 11/750 (Unix)	CRAY/XMP
PTN150	1.98	23.86	0.165
PTN660	22.85	297.93	1.402
SMBANK	2.64	21.50	0.177
BIGBANK	376.74	9100.00	58.896
GROUP1ac	218.46	1652.00	4.983
GROUP1ad	1320.82	10227.00	49.454
MARK3	24.87	204.43	2.131
Average	281.19 (17)	3075.25 (184)	16.744 (1)

Table 4 : Testing **NLPNETG** on different computer systems

The preceeding analyis for the Indianapolis sector was carried out on a VAX 11/750 minicomputer. The CPU time required to analyze the sector under different scenarios for risk and cost varied between 18 and 1200 sec. The VAX is a small computer by todays standards. To introduce additional elements in our model, and expand the analysis to cover the whole U.S. air-space will require access to a supercomputer. To demonstrate the ability of the network optimization algorithms to solve ultra-large problems in a matter of seconds refer to Table 4. We observe from this table that problems that take hours of CPU time on a VAX can be solved in less than a minute on a CRAY X-MP/24 vector supercomputer. This level of efficiency however can be achieved only if care is taken to streamline the software system for the vector environment of the CRAY. Refer to the paper by Zenios and Mulvey[14] for additional details.

7. Future Research

The network planning model OP holds promise for assisting air traffic controllers when the NASP becomes operational. In the meantime, research should be carried out on several fronts.

There are various OP modeling parameters which must be estimated, including the acceptable level of congestion in the sectors and the costs of delays. In addition, the airspace regions (R) and time periods

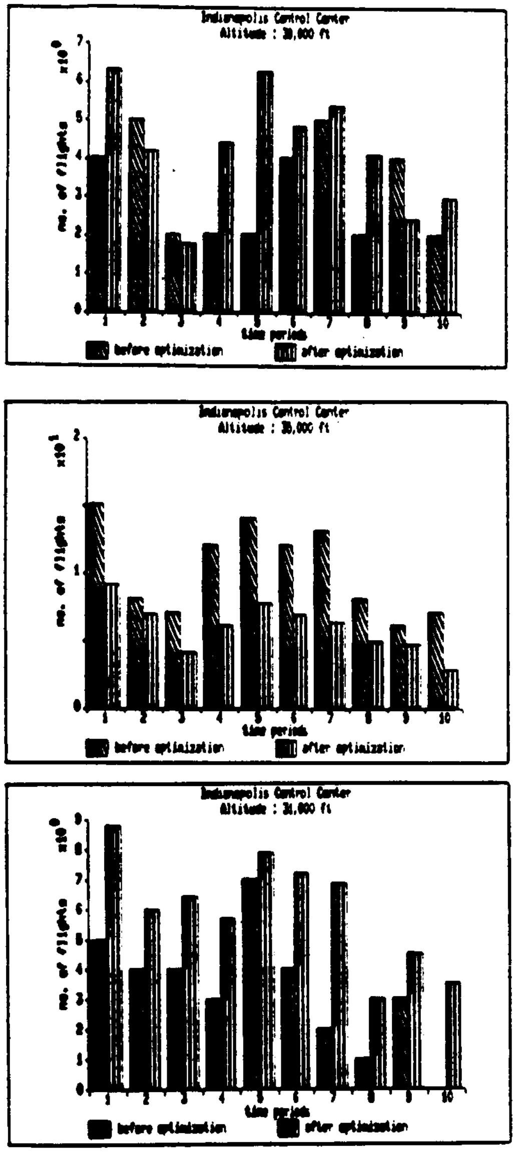

Figure 5 : Aircraft Congestion Before and After Optimization

(T) must be defined so that any boundary problems-- airplanes passing between regions-- are minimized and so that a practical-size system results. At the same time, the model must depict the real world to an acceptable level. Flight paths must be determined, including paths which are sparsely traveled, for each airport pair and aircraft type.

Nonlinear algorithms must be further studied; as new advances are made, they must be incorporated so that realistic-size examples will be solved. Remember that the OP model will be applied in a real-time environment. At present, two promising ideas are exact penalty methods and successive quadratic programming[5]. These algorithms will need to be further specialized for networks and for the computer architecture they are intended to run on. In conjuction, improved data storage schemes need to be discovered, schemes which are applicable to multi-processors.

The interface between the model and the human needs to be studied. Most large scale mathematical modeling systems do not include an extensive graphical display capability. This research, combining ideas from artificial intelligence and operations research, will become essential as the model and data become more complex.

We have assumed that data collection will be taken into account as part of the NASP. Much of the current air traffic system is automated, as witnessed by the automatic data collection efforts in the flow control project. Yet it is not unreasonable to expect that incompatibility and incompleteness of data will encountered. Any attempts to reduce this potential implementation problem will be helpful. It should be remembered that thousands of data items must be processed in a very short time.

References

1. *National Airspace System Plan,* U.S. Department of Transportation, Federal Aviation Administration, April, 1985.

2. D. P. Ahlfeld, R. S. Dembo, J. M. Mulvey, and S. A. Zenios, "Nonlinear Programming on Generalized Networks," Report EES-85-7, submitted for publication to *Transactions on Mathematical Software*, Princeton University, June 1985.

3. D. P. Bertsekas and D. El Baz, "Distributed Asynchronous Relaxation Methods for Convex Network Flow Problems," MIT Report LIDS-P-1417, October 1984.

4. M. Bielli, G. Calicchio, B. Nicolette, and S. Ricciavdelli, "The Air-traffic Flow Control Problem as an Application of Network Theory," *Computations and Operations Research*, vol. 9, no. 4, pp. 265-278, 1982.

5. M. C. Biggs, "On the Convergence of Some Constrained Minimization Algorithms Based on Recursive Quadratic Programming," *Journal of the Institute of Mathematics and Applications*, vol. 21, pp. 67-83, 1978.

6. G.R. Booth and C.M. Harvey, "Managing Air-traffic Delays by Mathematical Programming," *Proceedings of the Conference of Safety Issues in Air Traffic Systems Planning and Design*, Princeton University, Sep., 1983.

7. R. S. Dembo and J. G. Klincewicz, "Dealing with Degeneracy in Reduced Gradient Algorithms ," *Mathematical Programming* , vol. 31 , no. 3, pp. 357-363, March, 1985.

8. J. L. Helms, "Implementing the National Airspace System Plan," Safety Issues in Air Traffic Systems Planning and Design, Princeton University Conference, Sept. 1983.

9. J. M. Mulvey and S. Zenios, "Solving Large Scale Generalized Networks," *Journal of Information and Optimization Science*, vol. 6, pp. 95-112, 1985.

10. J. M. Mulvey, S. A. Zenios, and D. P. Ahlfeld, "Simplicial Decomposition for Convex Generalized Networks," Report EES-85-8, Princeton University, 1985.

11. J. M. Mulvey and S. A. Zenios, ''Integrated Risk/Cost Planning Models for the U.S. Air Traffic System,'' Report EES-85-9, Princeton University, June 1985.

12. A. Odoni and S. Endoh, ''A General Model for Predicting the Frequency of Air Conflicts,'' Safety Issues in Air Traffic Systems Planning and Design, Princeton University Conference #115, Sept. 1983.

13. S. A. Zenios and J. M. Mulvey, ''A Distributed Algorithm for Convex Network Optimization Problems,'' Report EES-85-10, Princeton University, Sept. 1985.

14. S.A. Zenios and J.M. Mulvey, ''Nonlinear Network Programming on Vector Supercomputers,'' Report EES-85-13, Princeton University, Feb., 1986.